

# NordTerm 2019

## Gæsteforelæsere og foreløbig liste af abstracts

---

### Gæsteforelæsere

**Andrejs Vasiljevs**, Tilde, Letland: Terminology in the age of AI – decline or revival?

**Frieda Steurs**, KU Leuven: Terminology 4.0: Enabling Knowledge Exchange and Transfer

**Sabine Kirchmeier**: Digital service i verdensklasse. Nye perspektiver for kunstig intelligens, sprogteknologi og terminologi i den offentlige sektor i Danmark

### Foreløbig liste over abstracts

#### Indholdsfortegnelse

Tommi Bernitz: Anvendelse af maskinoversættelse i Europa-Parlamentet: behov og muligheder for integration af terminologi...	2
Lotte Weilgaard Christensen, Bodil Nistrup Madsen: Terminologisk ontologi for 'indsatsledelse' på dansk...	3
Kristine Eide: Språkressurser i språkteknologien – termenes uerstattelighet og hvordan overtale offentlige etater til å ta ansvar for språkteknologi...	4
Lina Henriksen, Sussi Olsen, Claus Povlsen: eTranslation TermBank, terminologi på tværs af Danmark - og EU...	5
Brian Jacobsen: A Practical Example of the Use of Terms in Automated Classification...	6
Stine Jensen: Abstract on Nordic European Terminology Collaboration and Networking ...	7
Lars Bagøien Johnsen: Digitaliserede tekster for terminologisk analyse...	8
Susanne Lervad, Christian Gaubert: The Thread Project and Multilingual Textile Terminology...	9
Kerstin Lindmark, Julia Lautin, Niklas Jonsson, Mia Jensen: Scania Lexicon i verksamhetens verklighet och verktyg: terminologi i praktisk tillämpning...	10
Henrik Nilsson: Svårt med facilitering? – om terminologiska arbetsgrupper, konsensus och terminologens roll...	11
Henrik Nilsson: Lagringsvaror men inga lagervaror – om europeiska nationella termbanker och deras programvaror...	12
Niina Nissilä & Anita Nuopponen: Hur mycket språkteknologi behövs det i terminologiutbildningen?..	13
Gerd Sjögren & Mats Granström; Henrik Andersen: When Automation Requires a Semantic Interface: Some Examples...	14
Gunnhildur Stefánsdóttir: Developing, revising and using a terminology base – experience from the EEA Translation Centre in Iceland...	15
Mari Suhonen, Katri Seppälä: Digitalisering av offentliga tjänster – begreppsarbete som en del av utveckling och digitaliseringen av statsunderstödsverksamheten...	16
Hanne Erdman Thomsen: Relationer i Skatteforvaltningens Informationsmodeller...	17
Ágústa Þorbergsdóttir: Nye perspektiver og udfordringer for terminologi i en islandsk kontekst...	18
Marit Torseth, Jim Yang, Ole Våge: Høsting av digitaliseringens frukter: Hvordan begrepskataloger kan bidra til språkteknologi...	19

# Anvendelse af maskinoversættelse i Europa-Parlamentet: behov og muligheder for integration af terminologi

**Tommi Bernitz**

Europa-Parlamentet, GD TRAD, Den Danske Oversættelsesenhed

Luxembourg

[tommi.bernitz@europarl.europa.eu](mailto:tommi.bernitz@europarl.europa.eu)

## Maskinoversættelse i Europa-Parlamentet - eTranslation

Maskinoversættelse er almindeligt udbredt i dag blandt oversættere og vil sikkert vinde mere indpas i takt med, at maskinoversættelsesprogrammerne bliver bedre og mere tilgængelige for oversættere.

Europa-Parlamentet anvender ligesom andre EU-institutioner Europa-Kommissionens maskinoversættelsesprogram eTranslation til oversættelse af forskellige typer af dokumenter. Det dækker alle 24 EU-sprog med i alt 552 kombinationer. Alle oversættelser til og fra engelsk er direkte. De andre sprogpar bruger et system, hvor teksten først oversættes til engelsk og derfra til det endelige målsprog. eTranslation bruger Euramis-oversættelsehukommelser<sup>1</sup> som korpus. Siden 2018 har eTranslation anvendt neural maskinoversættelse.

## Behov og muligheder for integration af terminologi

Maskinoversættelse har bestemt ikke mindsket behovet for terminologi i oversættelsen. Selv om maskinoversættelse, herunder neural maskinoversættelse, generelt letter oversætternes arbejde, er der stadig behov for terminologi for at sikre præcisionen og korrektheden af den terminologi, der anvendes i oversættelserne, for at kunne levere kvalitetsfyldte oversættelser.

Det er derfor meget vigtigt, at der anvendes den samme terminologi i hele oversættelsen, og maskinoversættelsesprogrammerne er ikke altid i stand til at sikre dette. Der er derfor stadig brug for, at oversættere eller terminologer sikrer, at den korrekte terminologi anvendes.

En mulighed for direkte at integrere terminologi i maskinoversættelsen vil være at fodre og træne maskinoversættelsesprogrammer med terminologi fra IATE2-termbasen<sup>2</sup> for at få bedre terminologi. En anden mulighed vil være at bruge CAT-værktøjer sammen med termbaser, der oprettes ved hjælp af termekstraktion til at sikre mere konsekvent og korrekt terminologi. I Europa-Kommissionen eksperimenteres der med at finde nye måder til at integrere terminologi i maskinoversættelsen.

---

<sup>1</sup> Med over 1 mia. sætninger på de 24 officielle EU-sprog, der er produceret af EU-institutionernes oversættere.

---

<sup>2</sup> EU-institutionernes termbase IATE - Interaktiv Terminologi for Europa.

# Terminologisk ontologi for 'indsatsledelse' på dansk

**Lotte Weilgaard Christensen, Bodil Nistrup Madsen**

Syddansk Universitet, Copenhagen Business School  
Universitetsparken 1, DK-6000 Kolding, Dalgas Have 15, DK-2000 Frederiksberg  
[lotte@sdu.dk](mailto:lotte@sdu.dk); [bnm.msc@cbs.dk](mailto:bnm.msc@cbs.dk)

I forbindelse med en artikel i en antologi inden for 'disaster management' (Christensen & Madsen 2019) har vi udarbejdet en terminologisk ontologi for 'indsatsledelse' baseret på en alfabetisk ordliste fra Beredskabsstyrelsen (2018). Ordlisten indeholder 46 danske begreber med forklaringer. I artiklen beskriver vi principper for terminologisk arbejde og fokuserer på hvordan man ved at strukturere information ved hjælp af den terminologiske metode kan opnå en fælles forståelse blandt aktørerne fra forskellige beredskaber og hos borgerne i det hele taget.

Et første formål med dette bidrag er at redegøre yderligere for selve arbejdsprocessen ved oprettelse af terminologiske ontologier. Med den brugerinddragende metode kortsortering er det muligt at skabe sig et første overblik over relevante begreber og deres indbyrdes relationer for derefter at udarbejde begrebssystemer som input til de mere formelle terminologiske ontologier. Vi vil diskutere og redegøre for udfordringer i arbejdsprocessen og de foretagne valg. Udfordringerne har bestået i at begreberne i ordlisten ikke har været tilstrækkelige til at skabe en komplet ontologi. Der har fx manglet generelle begreber. Her har vi kunnet trække på FORVIR-topontologien. Derudover har der manglet ikke-leksikaliserede begreber. Endelig har vi konstateret polysemi i de øverste niveauer af ontologien.

I bidraget vil vi tillige beskrive og diskutere hvordan forklaringer i ordlisten kan omskrives til intensionale definitioner og formaliseres i form af træk-værdi-par.

I en perspektivering vil vi berøre hvordan terminologisk arbejde vil kunne (gen)anvendes i et grænseoverskridende samarbejde, som fundament for

klassifikationer og tesaurusser og som input til andre sprogteknologiske ressourcer/applikationer. Herunder vil vi relatere undersøgelsen til anbefalingerne fra Sprogteknologiudvalget i Danmark. Et perspektiv kunne her fx være automatisk indhentning og klassificering af informationer om katastrofer med henblik på hurtigere og bedre katastrofehandtering, evt. på tværs af grænser.

## Referencer

- Aasgaard, Marianne (2017): SPRÅKET I NØDETATER – forklare, ikke forvirre. In: Þorbergsdóttir, Ágústa (ed.): Forvaltning af fagsprog i samfundet, Hvem har ansvaret? Hvem tager ansvaret? REYKJAVÍK 9.-12. Juni 2015, NORDTERM 19, Island, 83-85.
- Beredskabsstyrelsen (2018): Retningslinjer for indsatsledelse. (<https://brs.dk/viden/publikationer/Documents/Retningslinjer%20for%20indsatsledelse%202018.pdf>).
- Christensen, Lotte Weilgaard & Bodil Nistrup Madsen (2019): A Danish Terminological Ontology of Incident Management in the Field of Disaster Management (*under udarbejdelse*).
- FORVIR: Topontologien <http://blog.cbs.dk/forvir/>
- Spencer, Donna (2009): Card sorting: Designing usable categories. Brooklyn, New York: Rosenfeld Media.
- Sprogteknologiudvalget i Danmark (2018): SPROGTEK2018 <http://sprogtek2018.dk>.

# Språkressurser i språkteknologien – termenes uerstattelighet og hvordan overtale offentlige etater til å ta ansvar for språkteknologi

**Kristine Eide**

Språkrådet

kristine.eide@sprakradet.no

I digitaliseringsprosessene i flere skandinaviske land, er begrepsarbeid en viktig premiss for at offentlige etater skal kommunisere med hverandre, og flerspråklige termlister er nødvendige når kommunikasjonen foregår over landegrenser. Offentlige etater må dermed ta ansvar for begrepsapparatet innen deres område.

Resultatet av dette begrepsarbeidet kan gjenbrukes til andre formål. En av de mest grunnleggende ressursene for automatisk oversettelse og mye annen språkteknologi er nettopp termlister.

Moderne språkteknologi fungerer foreløpig best på avgrensede fagområder. For at en talemåte skal forstå hva legene sier, må den føres med medisinske

uttrykk. En banks praterobot, må føres med bankterminologi. Jo mer spissede data en automatisk oversettelse arbeider med, jo bedre blir oversettelsen innenfor det samme avgrensede område. Derfor er for eksempel EUs system for automatisk oversettelse laget slik at man kan spesifisere innenfor hvilket område en tekst skal oversettes.

En termliste som er utarbeidet for et spesifikt område, kan dermed gjenbrukes til andre formål enn den opprinnelig var tiltenkt, vel og merke hvis den er tilrettelagt for gjenbruk. Dette må tas hensyn til i begrepsarbeid, og særlig offentlige etater må være klar over hvilken merverdi resultatet av terminologiarbeid kan tilføre både egen etat og samfunnet for øvrig.

# eTranslation TermBank, terminologi på tværs af Danmark - og EU

Lina Henriksen, Sussi Olsen, Bolette S. Pedersen, Claus Povlsen

Center for Sprogteknologi, NorS, Københavns Universitet

Emil Holms Kanal 2, 2300 København S

{linah, saolsen, bspedersen, cpovlsen}@hum.ku.dk

Projektet eTranslation TermBank startede i september 2017, løb i 1 ½ år og er altså netop afsluttet. Det var et EU-projekt under CEF-programmet – Connecting Europe Facility<sup>3</sup> – som investerer i og støtter opbygningen af digitale infrastrukturer der kan forbedre hverdagen for europæiske borgere. eTranslation TermBank havde til formål at bidrage til at styrke både kvaliteten og dækningsgraden af the Connecting Europe Facility Automated Translation platform (CEF eTranslation), en ny og forbedret udgave af EU's maskinoversættelsessystem, ved at forsyne systemet med terminologi på alle EU-sprog samt norsk og islandsk. CEF eTranslation er tilgængeligt for alle offentligt ansatte i EU-landene, og systemet vil på sigt blive tilgængeligt for alle europæiske borgere.

Projektet blev koordineret af Tilde fra Letland, som har arbejdet med lokalisering, maskinoversættelse og terminologi i mange år, og havde derudover otte deltagere fra syv lande, heriblandt CST på Københavns Universitet.

Helt konkret skulle hver partner i projektet identificere, og indsamle terminologiske ressourcer, fortrinsvis bilingvale eller multilingvale, med fokus på fagområderne sundhed, jura (forretningsjura) og forbrugerbeskyttelse for specifikt at styrke maskinoversættelse inden for nedenstående digitale service-infrastrukturer:

- ODR - online dispute resolution<sup>4</sup>
- e-Justice<sup>5</sup>
- eHealth

For at dække alle de nævnte sprog havde hver partner i projektet til opgave at indsamle terminologi for sit eget samt et eller flere andre sprog. De indsamlede ressourcer blev herefter processeret af Tilde og gjort klar til at kunne indgå i CEF eTranslation-systemet.

I vores oplæg vil vi fortælle lidt mere om baggrunden for projektet, om erfaringerne fra og resultaterne af indsamlingen i Danmark såvel som i det øvrige Europa, og informere om den fortsatte mulighed for at levere terminologi til CEF eTranslation-systemet via European Language Resource Coordination<sup>6</sup>.

## Referencer

- Auksoriūtė, A., Belogrīvs, I., Bielevivcienė, A., Blaudums, A., Bordāns, J., Borkowski, T., Borzovs, J., Braasch, A., Cauna, E., Dravniece, A., Dudlauskienė, N. & Galinski, C., Henriksen, L., Kaalep, H., Kalniņš, A., Kierzkowska, D., Kilgi, H., Kis, B., Liedskalniņš, A., Zabielaite, J. (2006). *Towards Consolidation of European Terminology Resources. Experience and Recommendations from EuroTermBank Project*. Tilde, Riga, Latvia.
- Gornostaja, T., Auksoriūtė, A., Dahlberg, S., Domeij, R., van Dorrestein, M., Hallberg, K., Henriksen, L., Kallas, J., Krek, S., Lagzdīņš, A., Lilles, K., Mitkevičienė, A., Olsen, S., Pedersen, BS, Pesliakaitė, E., Povlsen, C., Repar, A., Rozis, R., Sauberer, G., Thorbergdóttir, A., Vasiljevs, A., Vasiļevskis, A., Vaus, M. Zabarskaitė, J. (2018). *eTranslation TermBank: stimulating the collection of terminological resources for automated translation: Proceedings of the XVIII EURALEX International Congress*. 52. Ljubjana, Slovenien.

---

<sup>3</sup> <https://ec.europa.eu/inea/en/connecting-europe-facility>

<sup>4</sup> <https://ec.europa.eu/consumers/odr/main/index.cfm?ent=main.home2.show&lng=DA>

<sup>5</sup> <https://beta.e-justice.europa.eu/?action=home&plang=da>

---

<sup>6</sup> <http://www.lr-coordination.eu/>

# A Practical Example of the Use of Terms in Automated Classification

**Brian Jacobsen**

Chefkonsulent

Taxon ApS, Østergade 1, 1100 København K

Brian@Taxon.dk, +45 2390 5404

**Keywords:** Automated Classification, Process Automation, Metadata

An important part of automating processes is the ability to classify (or “tag”) texts according to one or more taxonomies.

The resulting metadata is vital in order to exchange, find and verify data between various systems, internal as well as external.

Terms in the sense of one or more coherent words are the very foundation of the automated classification system.

The presentation presents the automated classification system Taxon Classifier and the corresponding taxonomy tool TaxonHub. Together the two systems facilitate automated classification.

This will include a discussion about terms that combined defines a domain specific class, i.e. “super synonyms”, the use of statistics in combination with a rule engine in automated classification as well as practical examples of running systems like automated distribution of incoming mails using OS2KLE (the taxonomy used by all Danish municipalities) and access control in ESDH/ECM.

The goal of the presentation is to pave the way to a greater understanding of the important role terminology plays in automated classification and the automatization of processes in general.

# Abstract on Nordic European Terminology Collaboration and Networking

## Stine Jensen

Terminologist at TermCoord  
European Parliament  
Plt. du Kirchberg, 2929 Luxembourg  
Tél. +352 4300 21531 / +352 621 409 570  
stinekyung.jensen@europarl.europa.eu / stinekyungjensen@gmail.com

Terminology networking, expert involvement and collaboration is not a new thing within European terminology, and different language communities have already established language councils and networks for such purpose. For example: Danish (the Danish Terminology Network), Gaelic (the Expert Question Box), and Italian (REI - Rete per l'eccellenza dell'italiano istituzionale). Building upon this, the Terminology Coordination Unit of the European Parliament (TermCoord) last year made an effort to enforce this kind of networking with the further development of the [Expertise Network](#).

The objective of this collaborative terminology is to have the best possible quality in the database for European Terminology [IATE](#).

Since 2012, TermCoord has formalised terminology collaboration with terminology or language departments of universities and other interested parties: such as Translators Without Borders on terminology projects. A recent example of such collaboration is [Terminology without Borders](#), which combines expert involvement, expert validation and university collaboration within different fields: Health, fisheries, Education and women's rights.

The participating universities are predominantly from Southern European countries. As a result, TermCoord has, for the last year been working on balancing the representation of networks by strengthening the Nordic representation in collaboration with the Nordic language communities, such as Danish, Finnish and Latvian.

# Digitaliserte tekster for terminologisk analyse

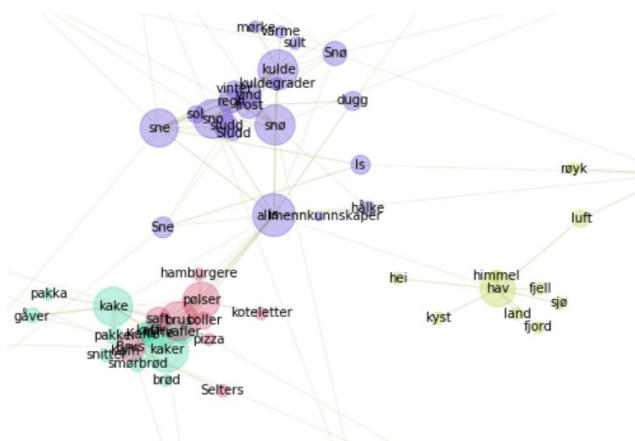
Lars Bagøien Johnsen  
Det Norske Nasjonalbibliotek  
Lars.johnsen@nb.no

## Abstract

Nasjonalbiblioteket i Norge har digitalisert hele bokarven, som er gjort tilgjengelig for terminologer og leksikografer. Vi skal spesielt se på hvordan sammenhenger og betydninger kan analyseres ved hjelp av kollokasjoner, og samforekomster i koordineringsmønstre.

**Keywords:** grafer, nettverk, betydning, cluster

I denne presentasjon vil vi gå gjennom den datamaskinnelle tilgangen til tekster ved Nasjonalbiblioteket i Norge. Til forskjell fra den konsumerbare lesning av tekster, som bare er tilgjengelig for norske IP-adresser, er den programmatisk tilgjengelig for alle. Den tilbyr konkordanser, n-grammer, ordgalakser, kollokasjoner og konkordanser.



Figur 1 Koordinasjonsgraf for ordet 'is'

Tekstene er gjort tilgjengelig på en slik måte at man kan analysere ord og deres betydninger gjennom kollokasjoner eller koordinasjonsgrafer. I Figur 1 er en koordinasjonsgraf for ordet 'is' som viser ord som clustrer seg sammen. Grafen skiller mellom flere betydninger, der de blå nodene utgjøre det stofflige, som i « is og snø », og de grønne « is » som dessert eller etterrett.

Dataene som er gjort tilgjengelig for leksikalske undersøkelser, utgjør ca 500 000 digitale bøker, og 800 000 aviser. For bøkene er alt av metadata tilgjengelig, som gjør det mulig å bygge korpus og undersøke ord innenfor bestemte grupper av tekster klassifisert i henhold til Dewey desimalklassifisering.

Bruken av metadata gjør det mulig å hente ut konkordanser, og hele setninger som belegg for bestemte betydninger.

For en leksikograf er dataene tilgjengelig i nettlesere, og benytter Jupyter Notebook som plattform.

Vi har data n-grammer, plotting over tid, og hvordan finne ut hvilke ord som har kraftige brudd i frekvens.

Hvordan henger ord sammen, og hvordan skille de forskjellige mønstrene fra hverandre.

## Språkressurser

Dataene er tilgjengelig fra den norske språkbanken, del av Nasjonalbiblioteket.



# The Thread Project and Multilingual Textile Terminology

**Susanne Lervad, Christian Gaubert**

Centre for Textile Research, University of Copenhagen

Karen Blixens Plads 8, DK-2300 København S

[Npf698@hum.ku.dk](mailto:Npf698@hum.ku.dk); [christian.gaubert@laposte.net](mailto:christian.gaubert@laposte.net)

Phone +4522110826

**Keywords:** textile, terminology, diachrony,

## **Textile terminology in the THREAD project - a multilingual approach**

Presentation of the THREAD project at the Centre for Textile Research (CTR) at the University of Copenhagen - a 3 year Innovation Foundation project in order to create a model for Education, Empowerment and Entrepreneurship for refugee women in Denmark.

The presentation reports on an internship of refugee women at CTR in 2017-2018 and the experiences and vocabulary results of the multilingual sessions on textile terminology. The terminology work includes

verbal and non-verbal representation of the concepts such as dress, photos, drawings, models in order to combine the different ways of representations of the basic concepts of the field.

Report of the internship of three refugee women from Eritrea and Turkey and multilingual sessions and documentation of textile vocabulary of knitting, tools and materials in Arabic, Danish, English and Tigrinya and thereby establishing a relationship with Eritrean refugees by exchanging knowledge on textile history.

# Scania Lexicon i verksamhetens verklighet och verktyg: terminologi i praktisk tillämpning

**Kerstin Lindmark, Julia Lautin, Niklas Jonsson, Mia Jensen**

Scania CV AB  
151 87 Södertälje

kerstin.lindmark@scania.com, julia.lautin@scania.com, niklas.jonsson@scania.com, mia.jensen@scania.com

**Keywords:** termdatabas, integration, kärnverksamhet

På Scania CV AB, som tillverkar lastbilar, bussar och industri- och marinmotorer, bedrivs sedan länge ett systematiskt terminologiarbete med syfte att skapa konsekvens och tydlighet, effektivisera och spara pengar. På Scanias intranät kan ca 16 000 personer läsa den terminologiska ordlistan Scania Lexicon direkt i webbläsaren. För att de eftersträvade effekterna ska kunna realiseras krävs dock mer än bara ett uppslagsverk: Användarna behöver få relevant språkstöd direkt i de system de arbetar i, så att det inte kostar extra att göra rätt. Användarnas roller, system och behov är vitt skilda, vilket ställer krav på ett mångfasetterat lexikon och systemstöd med möjlighet till anpassning och vidareutveckling i takt med att behoven förändras och utökas. En följd av detta är att

Scania Lexicon omfattar mer än bara termer, till exempel marknadsnamn, GUI-strängar och rubriker.

Scania Lexicons tillämpningar är inte i första hand språkteknologiska, men språkliga data slussas från termdatabasen till olika databassystem, översättningsmiljöer och content management-system, inklusive skribentmiljöer. Det krävs alltså viss databehandling för att vår terminologi ska bli tillgänglig och vägledande för våra användare i deras egna miljöer.

I det här föredraget kommer vi att med hjälp av ett antal use cases ge exempel på hur olika användare får tillgång till Scanias terminologi i sina egna verktyg, och på de utmaningar det innebär att få språkteknologiska verktyg integrerade i kärnverksamhetens system.

## **Svårt med facilitering? – om terminologiska arbetsgrupper, konsensus och terminologens roll**

**Henrik Nilsson**

terminolog

C.A.G Next

Kungsgatan 37, SE-111 56 Stockholm, Sverige

0046729614682

henrik.nilsson@cag.se

I några länder är ”terminolog” ett erkänt yrke, i andra mer okänt. Terminologiarbete utförs dock av många, oftast i grupp och ofta med en medverkande terminolog. Det speciella för terminologen är att denna person oftast själv inte har expertis på det fackområde som behandlas utan i stället fungerar, med modern terminologi, som en facilitator. Men också som en ”utvinnare”, en medlare och en katalysator. Hur går detta till och vad är det egentligen terminologen gör under och mellan möten? Vilka är de vanligaste problemen, lösningarna – och missuppfattningarna om denna yrkesroll? Hur kan terminologen fungera som katalysator – och vad händer om terminologen tar på sig eller får en ”diktatorroll”?

Detta föredrag kommer att, utifrån praktiska projekt, presentera (själv)reflektioner kring terminologens yrkesroll, funktion och beteende men också undersöka, utifrån teorier om konsensusskapande, arbetsprocesser i grupper där olika personer med varierande expertis medverkar. Föredraget kommer också att beröra vad terminologens expertis består av, hur terminologrollen har förändrats i takt med att terminologiarbetet vidgats i riktning mot ontologi- och kunskapshantering, modellering och mappning, och kort diskutera hur detta (bör) avspeglas i utbildning av terminologer.

# Lagringsvaror men inga lagervaror – om europeiska nationella termbanker och deras programvaror

**Henrik Nilsson**

ordförande

Europeiska terminologiföreningen (EAFT)/C.A.G Next

Kungsgatan 37, SE-111 56 Stockholm, Sverige

0046729614682

henrik.nilsson@cag.se

I samband med Nordterm 2019 ordnar Europeiska terminologiföreningen (EAFT) ett uppföljningsseminarium, TERMINTRA II, till det seminarium som hölls i Oslo 2012, om nationella termbanker. Sedan 2012 har fler länder skapat nationella termbanker eller utvecklat de befintliga, medan andra ännu inte har en enda samlande termbank på nationell nivå. Vad kan vi lära av varandra när det gäller just den här typen av termbank? Och åt vilket håll bör den utvecklas – mer eller mindre användarmedverkan, mer eller mindre normativt innehåll, enklare återbruk av innehållet?

De tekniska lösningar som valts för just nationella termbanker verkar variera stort, vilket i sig är intressant. I takt med att dagens terminologihanteringssystem (TMS:er) utvecklas borde de vara ett alternativ när stora termbanker skapas. Klara de inte de krav som just den nationella termbanken och dess vida målgrupp ställer? Eller handlar det helt enkelt om att det är enklare att få pengar för att skapa en egen programvara? Kan det handla om

konkurrens från andra, öppna lösningar av wikipertyp?

Dessa frågor kommer att diskuteras i detta föredrag som kommer att ta upp exempel på nationella termbanker i Europa och diskutera deras val av tekniska lösningar. Föredraget kommer också att beröra de uppdateringar som är på gång av relevanta ISO-standarder som rör termbanker och termbanksutveckling.

## Referenser

Nilsson, H. & Cuadrado, S. 2015. Software – Hard for National Termbanks? IITF Workshop. I: Lušický, V. & Budin, G. (red.), *Book of Abstracts of the 20<sup>th</sup> European Symposium on Languages for Special Purposes, 8–10 July 2015*, Vienna, Austria, Multilingualism in Specialized Communication: Challenges and Opportunities in the Digital Age, ISBN: 978-3-200-04186-8, <https://fedora.phaidra.univie.ac.at/fedora/get/o:406394/bdef:Content/download,2019-02-15>

# Hur mycket språkteknologi behövs det i terminologiutbildningen?

**Niina Nissilä & Anita Nuopponen**

Enheten för marknadsföring och kommunikation

Vasa universitet

niina.nissila@univaasa.fi, anita.nuopponen@univaasa.fi

**Nyckelord:** terminologiundervisning, terminologiförvaltning, språkteknologi

I vårt bidrag diskuterar vi förnyande av terminologiutbildningen vid Vasa universitet inom läroämnet kommunikationsvetenskaperna. Under de fyrtio åren som vi har haft terminologikurser har innehållet för det mesta fokuserat i de grundläggande terminologiska och begreppsanalytiska metoderna och principerna. På kandidatnivån (VIEK) riktar sig en grundkurs i terminologi till alla kommunikationsstuderande, på magisternivån har vi en kurs i begreppsanalys både i det multidisciplinära magisterprogrammet för kommunikation (VIMA) och i magisterprogrammet för teknisk kommunikation (TEVI). Därutöver har vi en specialisering i terminologi i det sistnämnda.

I detta föredrag fokuserar vi speciellt i utvecklingen av en kurs i terminologiarbete och -förvaltning som

kommer i fortsättningen att fungera som en fortsättnings- och specialiseringskurs på båda nivåerna (VIEK & TEVI). Målet är att studenterna ska känna igen terminologiska problem och kunna argumentera för behovet för terminologiarbete och -förvaltning i arbetslivet, och veta hur terminologins verktyg kan användas. På kursen behandlas även kriterierna för termbildning och -evaluering. Vidare ska studenterna bekanta sig med de grundläggande principerna för språkteknologiska applikationer som kan användas i terminologihantering och kunna utvärdera dem. Eftersom fokus i deras utbildning ligger i teknisk och fackkommunikation i stället för språk och översättning, letar vi efter ett svar på frågan vilka är de mest relevanta språkteknologiska element som ska ingå i kursen.

## When Automation Requires a Semantic Interface: Some Examples

**Gerd Sjögren<sup>1</sup> & Mats Granström<sup>2</sup>; Henrik Andersen<sup>3</sup>**

Director Business Partners<sup>1</sup>; Product Director<sup>2</sup>; Senior Partner<sup>3</sup>

Interverbum Technology<sup>12</sup>; Copenhagen Translation<sup>3</sup>

Box 1121, SE 16422 Kista, Sweden<sup>12</sup>; Hirsemarken 1, DK-3520 Danmark<sup>3</sup>

Stockholm/Sweden; Copenhagen/Denmark

{gerd.sjogren, mats.granstroem}@[interverbumtech.com](mailto:interverbumtech.com)<sup>12</sup>; [henrik.andersen@copenhagentranslation.dk](mailto:henrik.andersen@copenhagentranslation.dk)<sup>3</sup>

There are ambitious initiatives in sectors as diverse as Health, eCommerce and Process Industry to automate functions where semantic comprehension is a prerequisite for implementation. Without this comprehension, attempts to automate run the risk of becoming excessively time-consuming. The current presentation will focus on such cases. The first example is the Nordic Innovation-supported development project *Patient-Professional Communication Plugin (PPCP)*, where *free-text* extracts from an Electronic Health Record can be displayed, and difficult terms « translated » into everyday language, in real time. The interface allows the user to choose these « translations », originating from a termbase, in one or

several of the five major Nordic languages as well as English, Arabic, Somali and Tigrinya. A different example, using the same termbase, shows how someone writing a medical text can be prompted to use either a standardised medical term, or its everyday equivalent, rather than « professional jargon ».

Additional examples will focus on how to make it easy for a novice to tackle « controlled language », typically in the form of a drop-down menu, in situations where the individual does not recognise or fully understand the available alternatives. All the presented examples make use of a terminology tool working in the background.

# Developing, revising and using a terminology base – experience from the EEA Translation Centre in Iceland

**Gunnhildur Stefánsdóttir**

The Ministry for Foreign Affairs Translation Centre  
Rauðarárstígur 27, 105 Reykjavík, Iceland  
gunnhildur.stefansdottir@utn.stjr.is

**Keywords:** term base revision, terminology management, translation

The EEA Translation Centre in Iceland has been operated for almost three decades. Its main task is to translate the legal acts pertaining to the EEA Agreement. Integral to the translation environment is the management of terminology in the acts. The term base compiled by the translation centre team is one of the largest of its kind in Iceland, with over 80.000 entries at present. A significant factor in the translation is finding and determining Icelandic translations of the specialised terms.

In this presentation I will look into various aspects of the terminology work carried out in the centre. First, I will discuss how this task has developed since the work began. It gradually became clear that the term base needed revision, not just addition. There have been several phases of revision, and I

will describe the methodology used. Second, I will explore the needs of the translator as regards terms and their translation, and term entry structure and information. There are several factors pertinent to this examination. For instance, what makes a term entry reliable from the translators' perspective, how to structure data fields, and how terminologists can cater for the needs of translators. Thirdly, I will mention the issue of finding and acquiring appropriate terminology from external resources, both from field experts and valid linguistic resources on web.

It is critical that the term base helps the translator to choose and use consistent and correct terminology in the texts. To ensure that, an ongoing revision and quality control of the term base is necessary.

## Digitalisering av offentliga tjänster – begreppsarbete som en del av utveckling och digitaliseringen av statsunderstödsverksamheten

**Mari Suhonen, Katri Seppälä**

filosofie magister, terminolog; filosofie magister, direktör på TSK, terminolog  
Terminologicentralen TSK  
Runebergsgatan 4c B 20, FI-00100 Helsingfors  
tfn +358 9 27091063; tfn +358 9 27091062  
[mari.suhonen@tsk.fi](mailto:mari.suhonen@tsk.fi); [katri.seppala@tsk.fi](mailto:katri.seppala@tsk.fi)

**Nyckelord:** terminologiska metoder, begreppsanalys, terminologiska tjänster, terminologiska produkter, terminologiarbete, begreppsarbete, digitalisering av tjänster

I Finland har finansministeriet som resultat av en förutredning inlett ett omfattande projekt för utveckling och digitalisering av statsunderstödsverksamheten. Målet är att understödsverksamheten ska bli mer transparent och effektiv samt få större genomslag. Undervisnings- och kulturministeriet har satt igång med förberedelser som ska stödja projektet som finansministeriet kommer att tillsätta.

Projektet ska främja samarbetet mellan de statliga myndigheter som beviljar understöden. Det förverkligas genom att skapa en enhetlig och kundinriktad nationell modell för datasystem och gemensamma tjänster för myndigheternas gemensamma bruk. Detta kommer att minska mängden administrativt arbete i bidragsprocessen och den administrativa bördan även för dem som ansöker om understöd.

En gemensam verksamhetsmodell och ett enhetligt datasystem förutsätter dock att alla parter har en gemensam uppfattning om de centrala begreppen.

Därför bildade kultur- och undervisningsministeriet med stöd från Centret för IT-expertis CSC en arbetsgrupp med medlemmar från olika statliga myndigheter och startade begreppsarbetet med Terminologicentralen våren 2018.

Även om de statliga myndigheter som beviljar statsunderstöden följer statsunderstödslagen har deras verksamhet sina egna särdrag och det har uppstått olika tolkningar om de centrala begreppen. Dessutom föreskrivs vissa myndigheternas verksamhet närmare i andra författningar där det eventuellt används olika termer för samma begrepp. Därför är den gemensamma statsunderstödslagen inte en fullständig källa för begreppen och termer utan man måste reda ut hur de olika myndigheterna brukar begreppen och harmonisera begreppen så mycket som möjligt.

I vårt föredrag kommer vi att diskutera vad som behövs i begreppsarbetet som syftar till att skapa gemensamma riktlinjer, verksamhetsprocesser och digitala ärendehanteringssystem för flera myndigheter.



# Relationer i Skatteforvaltningens Informationsmodeller

**Hanne Erdman Thomsen**

Informationsarkitekt  
Skatteforvaltningen, Udviklings- og Forenklingsstyrelsen  
Østbanegade 123, København  
hanneerdman@gmail.dk

## Abstract

Skatteforvaltningens informationsmodeller i UML indeholder klasser med attributter og relationer mellem klasserne. Relationerne omfatter specialisering, som svarer til typerelationen, og uspecificerede relationer, som navngives ad hoc. På Handelshøjskolen CBS er der tidligere arbejdet med at opstille et inventar af relationer til brug i terminologiarbejde, og dette inventar har siden været anvendt i bl.a. Kriminalforsorgen. I dette indlæg vil jeg anvende relationsinventaret på eksempler fra Skatteforvaltningens informationsmodeller, og se om der er relationer i eksemplerne som ikke er dækket ind i inventaret.

**Keywords:** relationer, begrebsmodeller, informationsmodeller

## Semantiske relationer i ontologier

I forbindelse med projektet OntoQuery (Andreasen et al. 2002) arbejdede vi med semantiske relationer i ontologier (Madsen et al. 2002). I forlængelse af dette arbejde udarbejdede DANTERMcentret i samarbejde med Sundhedsdatastyrelsen et samlet inventar på 24 relationer omfattende typerelationen, fire forskellige del-helhedsrelationer, to temporale relationer og 17 associative relationer (Sundhedsdatastyrelsen (intet årstal)). De associative relationer er karakteriseret ved hvilke typer begreber de forbinder: aktiviteter, roller, lokationer samt den meget generelle type, entitet. Der er for eksempel 4 aktivitet-rolle-relationer: aktivitet-agent, aktivitet-patient, aktivitet-resultat og aktivitet-instrument. Alle relationerne navngives med et verbum (evt. med præposition) i hver retning. Således kaldes aktivitet-instrument-relationen henholdsvis *udføres ved hjælp af* og *anvendes ved*:

*Brændehugning udføres ved hjælp af en økse*

*En økse anvendes ved brændehugning*

DANTERMcentret har desuden anvendt relationsinventaret i forbindelse med praktisk begrebsafklaringsarbejde hos blandt andet Kriminalforsorgen i Danmark (Madsen, B.N. 2012).

## Skatteforvaltningens informationsmodeller

I Skatteforvaltningen udarbejdes der informationsmodeller i forbindelse med it-udviklingsarbejde, og i disse modeller navngives relationer ad hoc, uden en systematisk tilgang. Der er ikke tale om terminologiske begrebsmodeller, som de terminologiske begrebsrelationer er beregnet på, bl.a. er klasserne forsynet med attributter og der tilføjes kardinalitet på relationer.

## Ontologiske relationer i informationsmodeller

I mit indlæg vil jeg vise i hvilket omfang det er muligt at erstatte ad hoc-relationerne i Skatteforvaltningens informationsmodeller med de standardiserede relationer som anvendes i begrebsarbejdet hos Kriminalforsorgen.

## Referencer

- Andreasen, T., Jensen P.A., Fischer Nilsson, J., Paggio, P., Sandford Pedersen, B., and Thomsen H.E. (2002). OntoQuery: Ontology-based Querying of Texts. Trykt i: *Acquiring (and Using) Linguistic (and World) Knowledge for Information Access*. Jussi Karlgren, Pentti Kanerva & Björn Gambäck (ed.). Working Notes. March 25 - 27 2002, Stanford University. AAAI-2002 Spring Symposium Series, pp 28-31, Stanford.
- Madsen, B.N. (2012): *Vejledning for begrebsarbejde i Kriminalforsorgen*. Kriminalforsorgen og DANTERMcentret. Hentet 28-02-2019 på <http://kriminalforsorgen.iterm.dk/search/filer/H%C3%A5ndbog%20i%20begrebsarbejde%202012-09-18.pdf>
- Madsen, B.N., Sandford Pedersen, B., and Thomsen H.E. (2002). Semantic Relations in Content-based Querying Systems: a Research Presentation from the OntoQuery Project. Trykt i: Simov, K. & Kiryakov, A. (ed.): *Ontologies and Lexical Knowledge Bases: 1. International Workshop, OntoLex 2000, Sozopol, Bulgaria, September 8-10, 2000 – Proceedings*. Sofia; Onto Text Lab.

## Referencer til sproressourcer

- Sundhedsdatastyrelsen. (intet årstal). Begrebsbasen. Tilgået 28-02-2019 på <https://sundhedsdatastyrelsen.dk/da/rammer-og-retningslinjer/om-terminologi/nbs>.

## Nye perspektiver og udfordringer for terminologi i en islandsk kontekst

**Ágústa Þorbergsdóttir**

Stofnun Árna Magnússonar í íslenskum fræðum

redaktør

Laugavegi 13, 101 Reykjavík

354-5254440

agusta.thorbergisdottir@arnastofnun.is

Som bekendt er de såkaldte termudvalg en udbredt ordning i terminologiarbejde i Island. Det er oftest fageksperter, som står bag tilblivelsen af termlister og andet fagsprogsarbejde. Drivkraften for deres arbejde er i de fleste tilfælde baseret på deres interesse, og de deltager i et frivilligt arbejde. Blandt andet af denne grund kan termarbejdet være lidt tilfældigt, og det viser sig, at i den islandsk termbank, Íðorðabankinn, hvor de fleste termlister er publiceret, er det nemt at finde domæner hvor der mangler termlister.

Med projektet eTranslation TermBank (eTTB), som har til opgave at indsamle terminologiske ressourcer til at forbedre kvaliteten af EU's automatiske oversættelsessystem, er en ny kurs lagt i islandsk terminologi. Aldrig før er der blevet gjort en sådan målrettet indsats for at indsamle

terminologiske ressourcer i Island. I mit foredrag vil jeg diskutere resultaterne af projektet. Kom termlisterne i den islandsk termbank til gavn, og opnåedes der tilstrækkeligt gode resultater i de domæner, hvor der indtil nu har manglet islandsk termarbejde? Jeg vil også give et overblik over hvilke institutioner der har bidraget med terminologi til projektet, og nogle problemer som viste sig, for eksempel usikkerhed med opdateringer af versioner og ejendomsrettigheder.

Ifølge META-net-rapporten (2012) er digitale sprogressourcer på islandsk dårlige. Jeg vil argumentere for vigtigheden for denne indsamling af terminologiske ressourcer (eTTB) for små sprogsamfund som det islandsk, og som forhåbentligt kan anvendes videre i udvikling af sprogteknologiske værktøjer.

# Høsting av digitaliseringens frukter: Hvordan begrepskataloger kan bidra til språkteknologi

**Marit Torseth, Jim Yang, Ole Våge**

seniorrådgiver, seniorrådgiver, seniorrådgiver

Brønnøysundregistrene, Direktoratet for forvaltning og ikt (Difi), Språkrådet

Brønnøysundregistrene, Postboks 900, 8910 Brønnøysund; Difi, Postboks 1382 Vika, 0114 Oslo; Språkrådet, Postboks 1573 Vika, 0118 Oslo

marit.torseth@brreg.no, jim.yang@difi.no, ole.vage@sprakradet.no

+4797956521, +4799255998, +4745242401

Offentlig sektor i Norge har satt i gang et omfattende arbeid med å definere og beskrive begreper i forbindelse med digitalisering og deling av data i maskinlesbar form. Dette arbeidet kan gjøre det mulig å levere store mengder ett- og flerspråklige termressurser til språkteknologisk gjenbruk.

“Kun én gang” er et av prinsippene som ble fastslått i [Digital agenda for Norge](#). Det innebærer at offentlig sektor i mye større grad enn i dag skal gjenbruke data som allerede finnes. For blant annet å operasjonalisere “kun én gang” og gjenbruke data pågår det en stor satsing på informasjonsforvaltning i offentlig sektor. Derfor er det etablert en felles datakatalog med beskrivelser av datasett, begreper, API-er og informasjonsmodeller.

En av de viktigste forutsetningene for å gjenbruke data er at data er godt beskrevet, inkludert begrepene som dataene representerer. Om dataene skal gjenbrukes på tvers av virksomheter, må beskrivelsene av begrepene dessuten være standardiserte. Det er derfor fastsatt flere standarder

som beskriver hva som skal og kan tas med i begrepsbeskrivelser, samt hvordan begrepsbeskrivelser kan gjøres maskinelt tilgjengelig. Et viktig poeng er at standardene også legger til rette for flerspråklige beskrivelser av begreper.

Standardene utgjør en grunnmur som har gjort det mulig å utvikle en felles termportal for offentlig sektor, *Felles begrepskatalog*, som inneholder begreper fra flere store virksomheter. Felles begrepskatalog gjør det mulig å se likheter og forskjeller i begrepsdefinisjoner mellom ulike virksomheter, noe som er nødvendig når de samhandler med hverandre. Felles begrepskatalog gir også tilgang til et stort antall begreper fra offentlig forvaltning og gjør det mulig å høste dem til videre gjenbruk i språkteknologiske ressurser, f.eks. maskinoversettelse mellom norsk og engelsk og mellom bokmål og nynorsk. Gjennom Felles begrepskatalog vil språkbanker få tilgang til oppdaterte termressurser i sanntid.